





# Deep convolutional neural network-based scatterer density and resolution estimators in optical coherence tomography

THITIYA SEESAN,<sup>1,2</sup> IBRAHIM ABD EL-SADEK,<sup>1,3</sup>  PRADIPTA MUKHERJEE,<sup>1</sup> LIDA ZHU,<sup>1</sup> KENSUKE OIKAWA,<sup>1</sup> ARATA MIYAZAWA,<sup>1,4</sup> LARINA TZU-WEI SHEN,<sup>5</sup> SATOSHI MATSUSAKA,<sup>5</sup> PRATHAN BURANASIRI,<sup>2</sup>  SHUICHI MAKITA,<sup>1</sup>  AND YOSHIAKI YASUNO<sup>1,\*</sup> 

<sup>1</sup>Computational Optics Group, University of Tsukuba, Tsukuba, Ibaraki, Japan

<sup>2</sup>Department of Physics, School of Science, King Mongkut's Institute of Technology Ladkrabang, Ladkrabang, Bangkok, Thailand

<sup>3</sup>Department of Physics, Faculty of Science, Damietta University, New Damietta City, Damietta, Egypt

<sup>4</sup>Sky Technology Inc., Tsukuba, Ibaraki, Japan

<sup>5</sup>Clinical Research and Regional Innovation, Faculty of Medicine, University of Tsukuba, Ibaraki, Japan

\*yasuno@optlab2.bk.tsukuba.ac.jp

<https://optics.bk.tsukuba.ac.jp/COG/>

**Abstract:** We present deep convolutional neural network (DCNN)-based estimators of the tissue scatterer density (SD), lateral and axial resolutions, signal-to-noise ratio (SNR), and effective number of scatterers (ENS, the number of scatterers within a resolution volume). The estimators analyze the speckle pattern of an optical coherence tomography (OCT) image in estimating these parameters. The DCNN is trained by a large number (1,280,000) of image patches that are fully numerically generated in OCT imaging simulation. Numerical and experimental validations were performed. The numerical validation shows good estimation accuracy as the root mean square errors were 0.23%, 3.65%, 3.58%, 3.79%, and 6.15% for SD, lateral and axial resolutions, SNR, and ENS, respectively. The experimental validation using scattering phantoms (Intralipid emulsion) shows reasonable estimations. Namely, the estimated SDs were proportional to the Intralipid concentrations, and the average estimation errors of lateral and axial resolutions were 1.36% and 0.68%, respectively. The scatterer density estimator was also applied to an *in vitro* tumor cell spheroid, and a reduction in the scatterer density during cell necrosis was found.

© 2021 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

## 1. Introduction

Speckle is an inevitable phenomenon in the image formation of optical coherence tomography (OCT) [1]. The speckle is an information carrier that conveys the properties of tissues, such as the sub-resolution distribution of the scatterers [2] and sub-resolution translation [3,4]. In addition, speckle carries properties of the imaging system itself, such as the optical resolution [1]. We refer to these properties as “fundamental parameters” in this paper. This information-carrying nature of the speckle motivates us to measure/estimate the fundamental parameters through the speckle patterns of an OCT image.

However, the relationship between these parameters and the speckle pattern is complicated, and it is thus not straightforward to estimate these parameters from the OCT image. This complexity comes from the image formation process. In principle, OCT imaging processes can be understood as a combination of two sequential processes. The first process is a forward process in which the tissue property is encoded into an interference signal. The second process is the backward process in which the interference signal is processed to form an OCT image, and the OCT

image is then analyzed to estimate the fundamental parameters. The principle is exemplified by attenuation coefficient estimation [5–7], signal intensity estimators [8,9], and polarization parameter estimations of polarization sensitive OCT [10–14].

In general, the forward process is definitive and can be easily modeled, and also can be relatively easily numerically simulated [15–17]. On the other hand, the backward process is complicated and hard to be performed. Despite its ease of simulation, the forward process is a physical process and does not always need to be simulated. In contrast, despite the difficulty, the backward process should be numerically performed for accurate and quantitative OCT measurement.

Our fundamental idea is to use a deep convolutional neural network (DCNN) to solve the complicated backward process. In recent years, deep learning [18,19] has been demonstrated to be a powerful tool in the field of computer vision. Its application in OCT, such as in image segmentation [20–23], classification [24,25], and denoising [26,27] has been successful. The DCNN is one of the most established realizations of deep learning, which automatically learns a hierarchy of increasingly complex features related to the training data sets [28–30]. The DCNN is a highly nonlinear method and can therefore potentially solve the backward process.

In our approach, the DCNN takes a local (on the order of tens of micrometers) speckle pattern of an OCT image as an input and outputs fundamental parameters, including the scatterer density and optical resolution. Here the scatterer density denotes the number of scatterers per unit volume. However, the DCNN should be trained using a huge training data set, and it is not realistic to acquire such a huge data set in experiments. We here solve this issue to perform simple numerical simulations of the forward process. By employing a simple model of the OCT imaging, a huge number of local speckle patterns are created for several predefined resolutions and scatterer densities; i.e., the fundamental parameters. These sets of parameters and the generated speckle patterns are used to train the DCNN. Our strategy is therefore summarized as a backward process (i.e., estimation of parameters from a local speckle pattern) performed by the DCNN and the training of the DCNN using a fully numerically generated data set that is generated by simulating the forward process. These estimators are validated by numerically generated OCT images, experimentally obtained scattering phantom images, and *in vitro* tumor spheroid images.

## 2. Method

### 2.1. Three-dimensional speckle pattern generator

We generate small volumetric OCT speckle patterns in a simulation based on a simple Fourier imaging model. A three-dimensional (3D) scatterer distribution map,  $S(x, y, z)$ , is first generated stochastically to generate the speckle pattern. Here  $x$  and  $y$  are lateral positions and  $z$  is the depth position. The 3D scatterer distribution map is a complex 3D numerical array with  $128 \times 128 \times 128$  pixels, where the pixels with a scatterer (i.e., scatterer pixels) have an amplitude of unity and a random phase whereas the pixels without scatterer have zero values. The scatterer pixels are randomly selected such that the expectation of the scatterer density of the distribution map is a particular set value. The random phases of the scatterer pixels represent the randomly distributed sub-wavelength position of the scatterers.

This 3D scatterer distribution map is then numerically convolved with a 3D complex point spread function (PSF),  $\text{PSF}(x, y, z)$ , to yield a simulated complex OCT signal  $G(x, y, z)$  as

$$G(x, y, z) \equiv S(x, y, z) * \text{PSF}(x, y, z), \quad (1)$$

where  $G(x, y, z)$  is the 3D speckle pattern and  $*$  indicates the 3D convolution. The convolution is computed by a Fast-Fourier-Transform-based method. The PSF is numerically generated as a 3D Gaussian function with the same physical dimension and pixel numbers as the scatterer density

map,

$$\text{PSF}(x, y, z) \propto \exp - \left[ 4 \left( \frac{x}{\Delta x} \right)^2 + 4 \left( \frac{y}{\Delta y} \right)^2 + 2 \ln 2 \left( \frac{z}{\Delta z} \right)^2 \right], \quad (2)$$

where  $\Delta x$  and  $\Delta y$  are the lateral resolutions defined as the  $1/e^2$ -width of the squared amplitude of the PSF.  $\Delta z$  is the axial resolution defined as the full-width-half-maximum (FWHM) of the squared amplitude of the PSF. The lateral and axial resolutions are randomly selected, but two lateral resolutions  $\Delta x$  and  $\Delta y$  are selected to be identical. The axial and lateral resolutions are ones of our estimation target. Additionally, the magnitude of the PSF is randomly selected to simulate the uncontrolled probe beam power and unknown scattering intensity.

After computing the convolution [Eq. (1)], the 3D numerical field was down-sampled from  $128 \times 128 \times 128$  pixels to  $16 \times 16 \times 16$  pixels as keeping the original physical field size. After the down-sampling, we add complex Gaussian noise to each pixel to achieve a randomly but specifically selected signal-to-noise ratio (SNR) then taking the squared intensity of the 3D OCT speckle pattern.

The generated 3D speckle pattern covers a  $31.2\text{-}\mu\text{m}$  ( $x$ )  $\times$   $31.2\text{-}\mu\text{m}$  ( $y$ )  $\times$   $115.8\text{-}\mu\text{m}$  ( $z$ ) volume with  $16 \times 16 \times 16$  pixels. So, the physical size of a pixel is  $1.95\text{ }\mu\text{m}$  ( $x$ )  $\times$   $1.95\text{ }\mu\text{m}$  ( $y$ )  $\times$   $7.24\text{ }\mu\text{m}$  ( $z$ ), while that of the original  $128 \times 128 \times 128$ -pixel field is  $0.24\text{ }\mu\text{m}$  ( $x$ )  $\times$   $0.24\text{ }\mu\text{m}$  ( $y$ )  $\times$   $0.90\text{ }\mu\text{m}$  ( $z$ ). The pixel separations of the down-sampled speckle field are selected to be identical to those in our experiment described in Section 2.4. The inputs for the training of the DCNN-based estimator described in the next section are 2D cross-sectional OCT images, which are the  $x$ - $z$  cross-sections of the 3D numerically generated speckle pattern.

## 2.2. DCNN-based parameter estimator

### 2.2.1. Speckle data set

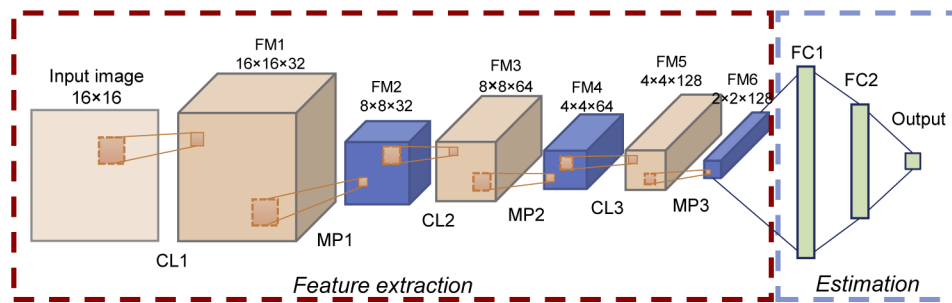
The data set for DCNN training is generated by the 3D speckle pattern generator described in the previous section. The true parameters are randomly selected with specific ranges of 3 to  $30\text{ }\mu\text{m}$  for the axial and lateral resolutions, 0 to 50 dB for SNR, and 0 to  $0.0645\text{ scatterers}/\mu\text{m}^3$  (0 to  $1.775\text{ scatterers}/\text{pixel}^3$ ) for the scatterer density. The maximum scatterer density corresponds to effective numbers of scatterers (ENS) of 33.7 and 76.7 for high- and low-NA objectives, respectively. Here the ENS is the number of scatterers within a 3D resolution volume, and see Section 2.4 for the details of the high- and low-NA objectives. The speckle patterns are normalized by intensity before being input to the DCNN, and the PSF magnitude is thus not a significant parameter. Each datum in the data set is the combination of the speckle pattern and the set fundamental parameters.

The training data set consists of 1,280,000 2D cross-sectional speckle patterns extracted from 80,000 3D speckle patterns. The validation data set, which is used to compute the loss during the training, comprises 160,000 2D speckle patterns extracted from 10,000 3D speckle patterns.

### 2.2.2. Network architecture and training

Our DCNN model architecture is shown in Fig. 1. The DCNN is input with the 2D cross-sectional speckle patterns, and it automatically learns the local features of the speckle pattern and estimates the underlying parameters. The whole network is subdivided into two parts, namely the feature extraction and the estimation parts.

The input to the feature extraction part is a 2D cross-sectional speckle image with a size of  $16 \times 16$  pixels [input image in Fig. 1]. The input image intensity is normalized to a  $[0, 1]$  range before being input. The first convolution layer (CL1) convolves the image with 32 different filters having a kernel size of  $2 \times 2$  to obtain a feature map (FM1;  $16 \times 16$  pixels  $\times$  32 filters). Here, the rectified linear unit (ReLU) is used as an activation function. A max-pooling layer (MP1) with a stride of 2 is applied after the convolution, which fuses nearby spatial information to reduce



**Fig. 1.** Architecture of the network used for parameter estimation. CL; convolution layer, FM; feature map, MP; max-pooling layer, and FC; fully connected layer. The numbers above the boxes are the sizes of the boxes as pixel  $\times$  pixel  $\times$  filters.

the size of the feature map. The second convolution layer (CL2) produces the third feature map (FM3;  $8 \times 8$  pixels  $\times$  64 filters) using  $2 \times 2$ -pixel kernels and ReLU activation. After applying a max-pooling layer with a stride of two (MP2), the third convolutional layer (CL3) generates the fifth feature map (FM5;  $4 \times 4$  pixels  $\times$  128 filters). Finally, the third max-pooling layer (MP3) makes the final feature map (FM6;  $2 \times 2$  pixels  $\times$  128 filters), and feeds the subsequent estimation part.

The estimation part consists of two fully connected layers of FC1 and FC2 which consist of 512 and 256 neurons, respectively. The final output layer consists of a single neuron and its output is an estimated fundamental parameter.

We train our neural network models to minimize the mean squared error (MSE) between ground-truth parameters and network outputs. The network is trained using an Adam optimizer [31] with a learning rate of  $10^{-4}$ . The batch size of the training was 32. Five independent networks with the same network architecture are trained for five fundamental parameters, namely the scatterer density, the lateral and axial resolutions, SNR, and ENS. It is noteworthy that the SNR is randomized and the signal strength is normalized, and hence, a DCNN might not use this information to estimate the scatterer density and ENS.

The models are constructed in Python 3.7 with Keras 2.3.1 based on the TensorFlow back end. The training of each network with a graphical processing unit (GPU; Nvidia GTX1080) takes about 5-6 min for each epoch. The training is terminated if the loss has not decreased for 10 continuous epochs, and the weights with the minimum loss are selected for the estimator. The minimum loss is obtained at the 30th, 28th, 27th, 15th and 28th epochs for the scatterer density, lateral and axial resolutions, SNR, and ENS, respectively.

### 2.3. Validation methods

The DCNN-based estimators are validated using numerically generated OCT signals, experimentally measured OCT images of the scattering phantom and an *in vitro* spheroid sample. The details are as following.

#### 2.3.1. Numerical validation

We evaluated the estimation performances of the DCNN-based estimators using 100 numerically generated OCT speckle patterns. The speckle patterns are generated with the same method used to generate the training data set, but they are not included in the training or validation data set. Each speckle pattern was extracted from a different simulated volume. So, all speckle patterns are independent of each other and have different fundamental parameters. Here, the evaluation performances for the scatterer density, ENS, the lateral and axial resolutions, and

SNR are evaluated. The true values are known for this evaluation, and we can thus obtain the estimation accuracies.

### 2.3.2. Validation by scattering phantom measurement

The DCNN-based estimators were validated experimentally by measuring scattering phantoms; the Intralipid emulsion. In the field of biomedical optics, 20% Intralipid emulsion (IL-20; I141, Sigma-Aldrich) is frequently used as an optical phantom to mimic the optical properties of tissue [32]. Six different dilutions were made by mixing the IL-20 with purified water to obtain concentrations of 1%, 2%, 4%, 6%, 8%, and 10%. Each dilution was then placed in a Petri dish. Black mending tape was attached at the bottom of the Petri dish to avoid strong specular reflection from the glass surface. Three samples were made for each concentration, i.e., 18 samples were made in total.

The scattering of light by Intralipid emulsion is due to the particles of soybean oil and egg lecithin, and the scattering volume concentration  $\sigma$  of IL-20 is known to be 22.7%(v/v) [33–35]. Hence, the scattering volume concentrations of the 1%, 2%, 4%, 6%, 8%, and 10% dilutions are 0.227, 0.454, 0.908, 1.362, 1.816, and 2.27%(v/v), respectively.

### 2.3.3. Human breast cancer spheroid

The scatterer density estimator was also applied to an *in vitro* sample. The sample was a human breast adenocarcinoma spheroid made from MCF7 cell line. The scatterers in the cell are cell nuclei and cell organelles, and we thus anticipated that the scatterer density was strongly related to tissue functions.

After 15-day cultivation, spheroids with a size of a few hundred micrometers had formed. The spheroid was extracted from the culturing environment and placed in a room-temperature culture medium without CO<sub>2</sub> supply. The cell might have been gradually dying because of a lack of nutrients. We performed OCT measurement every 2 hours up to 28 hours. Note that this experiment was originally performed for the study described in Ref. [36], and the same data set was used in the present study.

## 2.4. OCT setup and measurement protocol

Swept-source OCT with a 1.310  $\mu\text{m}$  wavelength probe beam and a measurement speed of 50,000 A-lines/s was conducted in the experiments [37]. Although the system was polarization-sensitive, we did not use its polarization functionality. The OCT image used in this study is the coherence composition of multiple polarization channels (Section 2.3.1 of Ref. [37] and Section 3.8 of Ref. [38]), which is almost equivalent to a standard OCT image.

Two objectives were used in this study. One has an effective focal length of 18 mm (LSM02, Thorlabs Inc., NJ). The beam diameter incident on the objective is 3.49 mm, and the effective numerical aperture (NA) was thus 0.097. This resulted in a diffraction-limit lateral resolution (spot size) [39] of 8.6  $\mu\text{m}$ , whereas the lateral resolution numerically simulated by OpticStudio (Zemax) is 8.9  $\mu\text{m}$ . We refer to this objective as a high-NA objective. The other objective, a low-NA objective (LSM03, Thorlabs Inc., NJ) had an effective focal length of 36 mm. The effective NA of 0.048 results in a diffraction-limit lateral resolution is 17.2  $\mu\text{m}$ , whereas the numerically simulated resolution is 18.1  $\mu\text{m}$ .

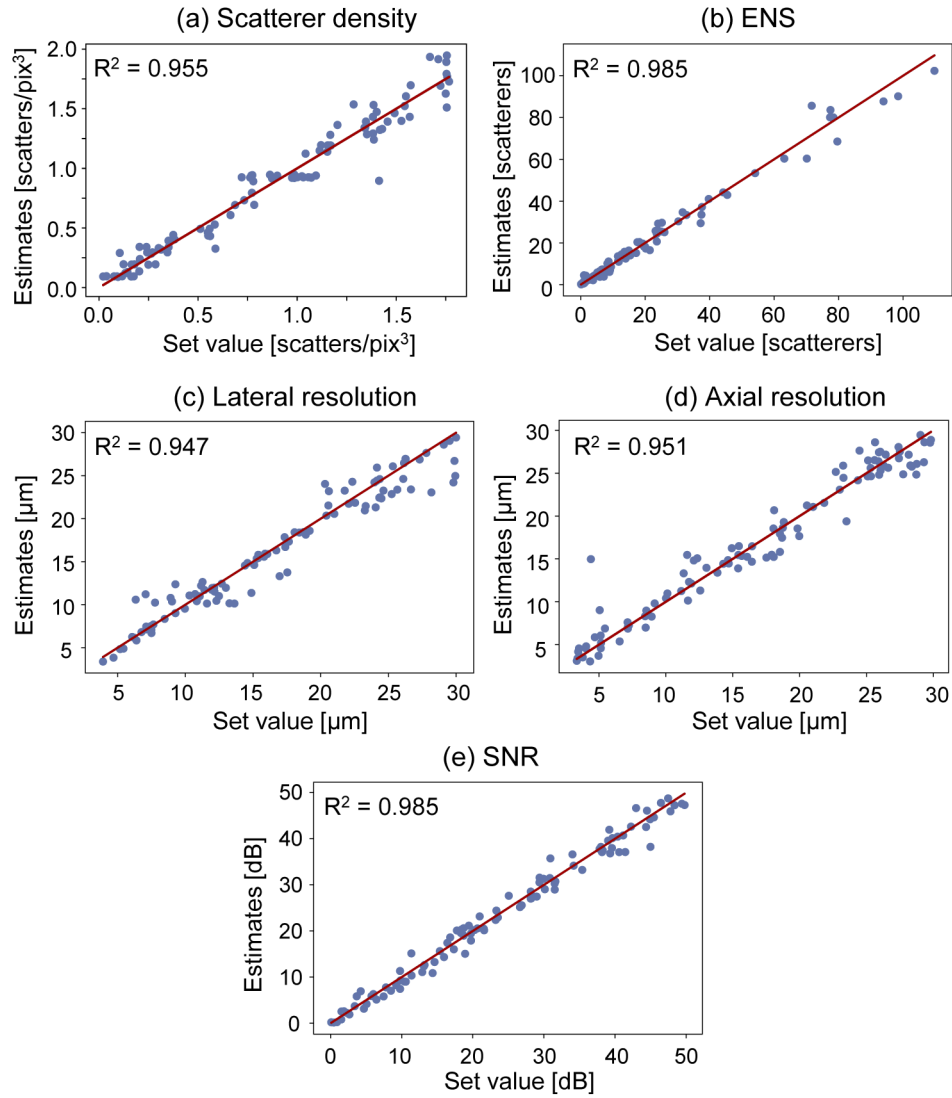
Each B-scan comprises 512 A-lines covering a lateral scanning range of 1 mm, and the lateral pixel separation is 1.95  $\mu\text{m}$ . The axial resolution of our system is 14.1  $\mu\text{m}$  and the axial pixel separation is 7.24  $\mu\text{m}$  in tissue. These axial and lateral pixel separations are identical to those of the numerically generated speckle patterns (Section 2.1).

The measurements were performed for several values of probe beam attenuation, namely 0 dB, −5.4 dB, and −11.76 dB (for round trip). The probe power on the sample was 12 mW with 0-dB attenuation.

### 3. Results

#### 3.1. Numerical validation

All DCNN-based parameter estimators show high consistency between the set and estimated parameters as shown in Fig. 2. Figure 2(a)-(e) are for the scatterer density, ENS, lateral and axial resolutions, and SNR, respectively. The horizontal axes represent the set parameters whereas the vertical axes show the estimates.



**Fig. 2.** Numerical validation of the parameter estimations of scatterer density (a), ENS (b), lateral resolution (c), axial resolution (d), and SNR (e). The horizontal and vertical axes give the set (true) values and the estimates, respectively. The red lines show the perfect estimation.

The root mean square errors (RMSEs) of the estimations are 0.113 scatters/ $\mu\text{m}^3$  for the scatterer density, 3.06 scatters for the ENS, 1.78  $\mu\text{m}$  for the lateral, 1.88  $\mu\text{m}$  for the axial resolutions, and 1.82 dB for the SNR. These RMSEs correspond to 0.23%, 6.15%, 3.58%, 3.79%,

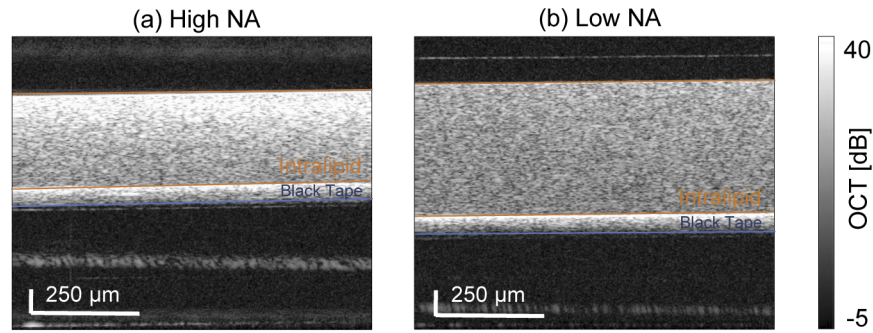


and 3.65% of the estimation ranges, respectively. Here, the estimation ranges are the ranges of each parameter in the speckle generation.  $R^2$  values of a model “estimate = true value” are 0.955 for the scatterer density, 0.985 for the ENS, 0.947 for the lateral resolution, 0.951 for the axial resolution, and 0.985 for the SNR. The estimators therefore give the reasonable estimates of the parameters.

### 3.2. Scattering phantom

#### 3.2.1. Scatterer density estimation

Figure 3 shows the example OCT B-scans of the scattering phantom with the concentration of 0.227%(v/v), where Figs. 3(a) and 3(b) are obtained with the high- and low-NA objectives, respectively. We process these images using the scatterer density estimator having a sliding window of  $16 \times 16$  pixels and computed the average scatterer density within the Intralipid region. Figure 4 shows the average scatterer density at each scatterer concentrations for each of three probe-beam attenuations. Figures 4(a) and 4(b) show results for the high-NA and low-NA objectives, respectively. Three phantoms were measured for each concentration, and the error bars in the plots represent the standard deviations among the three measurements of the three samples.



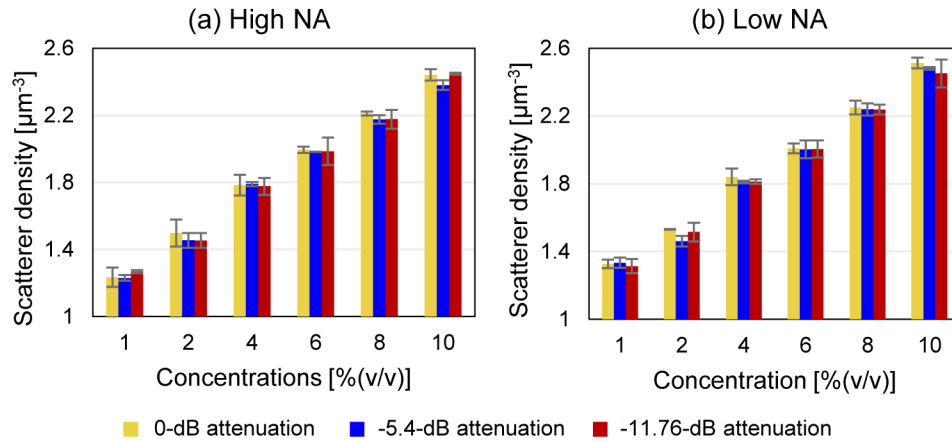
**Fig. 3.** Example OCT images of a scattering phantom with a concentration of 0.227%(v/v). (a) and (b) are images obtained with the high- and low-NA objectives, respectively.

It is found that the results are consistent for the high- and low-NA objectives. It is also noteworthy that the estimations are not affected by the probe beam attenuation, i.e., the probe power. This suggests that the estimator does not estimate the scatterer density from the OCT signal strength.

The estimated scatterer densities have a linear relationship with the concentration. The average slopes of the three measurements (three phantoms) are 0.5622 (for  $-0$  dB attenuation), 0.5436 ( $-5.4$  dB), and 0.5577 ( $-11.76$  dB)  $\mu\text{m}^{-3}/\%(v/v)$  for the high-NA objective, and 0.5560 ( $-0$  dB), 0.5587 ( $-5.4$  dB), and 0.5415 ( $-11.76$  dB)  $\mu\text{m}^{-3}/\%(v/v)$  for the low-NA objective.

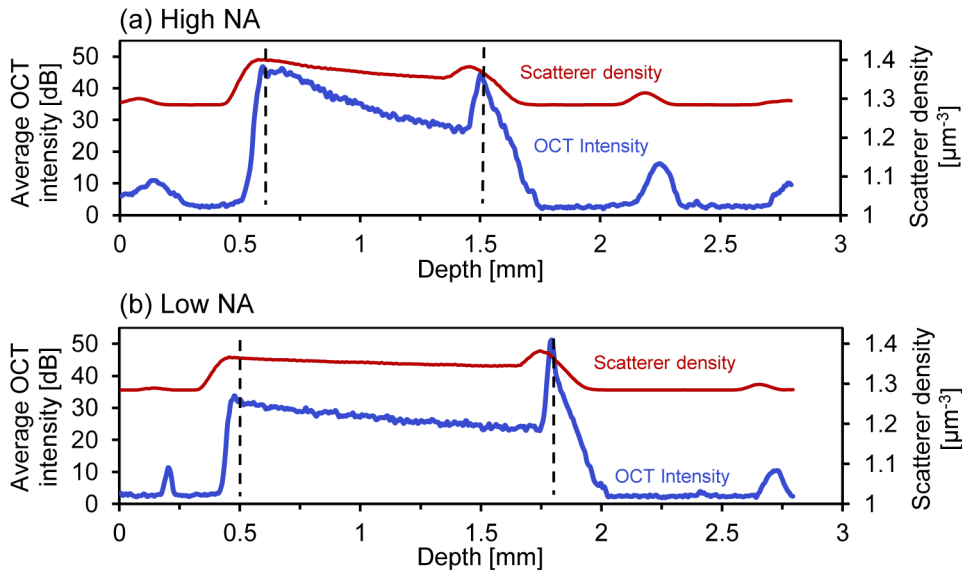
We can compute the scatterer size from these slopes as discussed in detail in Section 4.1. The intercepts of plots are found not to be zero, although the ideal intercept is zero. The intercepts are 1.23 (for  $-0$  dB attenuation), 1.20 ( $-5.4$  dB), and 1.20 ( $-11.76$  dB)  $\mu\text{m}^{-3}$  for the high-NA objective, and 1.26 ( $-0$  dB), 1.23 ( $-5.4$  dB), and 1.25 ( $-11.76$  dB)  $\mu\text{m}^{-3}$  for the low-NA objective. This issue is discussed in Section 4.4.

To evaluate the depth dependency of the scatterer density estimation, the estimated scatterer densities are averaged along the lateral direction (red lines in Fig. 5). Here, the data correspond to those in Fig. 3 and the concentration is 0.227%(v/v). The blue lines indicate the average OCT intensity and the region between the dashed lines is the region of the Intralipid. The strong OCT



**Fig. 4.** Scatterer density estimation of the Intralipid phantoms with several volume concentrations. (a) and (b) are the results obtained with the high- and low-NA objectives, respectively, and the color represent different attenuation of the probe beam. The error bars depict the standard deviations for three measurements of three phantoms.

signals at just beneath the Intralipid region (dash line) are from black mending tape placed at the bottom of the Petri dish.



**Fig. 5.** Depth profiles (blue lines) and the estimated scatterer density (red lines) of the phantom shown in Fig. 3.

It is found that the estimated scatter densities are almost constant along the depth, and they are consistent among the NAs. It is also found that the estimated scatterer densities are independent of the average OCT intensity. The mean scatterer densities in the Intralipid region are  $1.37 \text{ scatterers}/\mu\text{m}^3$  and  $1.35 \text{ scatterers}/\mu\text{m}^3$  for the high- and low-NA objectives, respectively, i.e., the difference is only 1.5%.

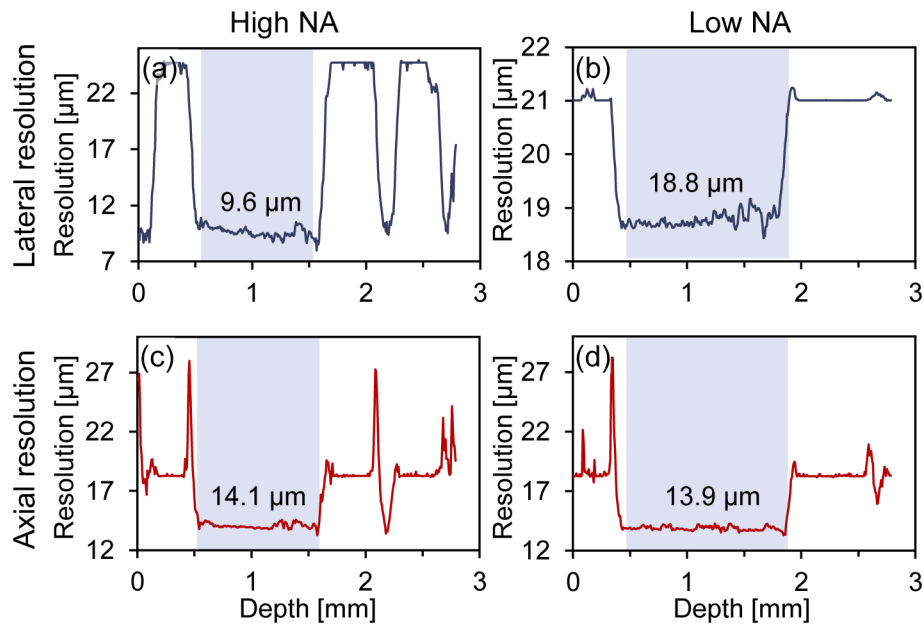
Attenuation coefficients or the signal attenuation rate along the depth are frequently used to evaluate the scattering property of a sample. The attenuation rates of these data are computed



through linear regression of the OCT intensity as  $-24.8$  dB/mm (for the high-NA) and  $-6.8$  dB/mm (for the low-NA). This relatively large inconsistency (8%) is accounted in part by the different focus positions and the different depth-of-foci of the two measurements.

### 3.2.2. Resolution estimation

Figure 6 shows the resolution estimation results of the scattering phantom. The data correspond to those of Fig. 3. The blue regions indicate the region of Intralipid.



**Fig. 6.** Resolution estimates of scattering phantoms (0.227%(v/v)). The first and second columns are for the high- and low-NA objectives, and the first and second rows show the lateral and axial resolutions, respectively. The blue background indicates the Intralipid region.

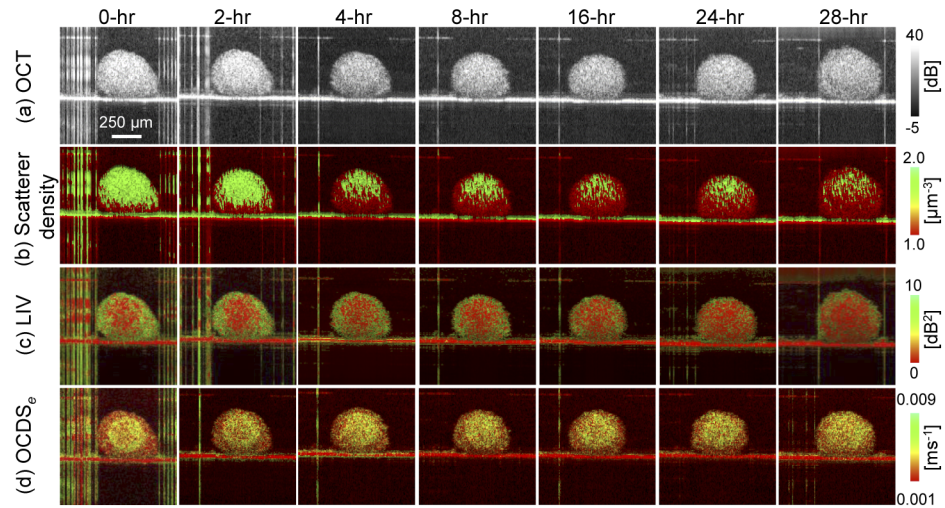
The average estimates of the lateral resolution over the Intralipid region were  $9.6 \mu\text{m}$  and  $18.8 \mu\text{m}$  for the high- and low-NA objectives, respectively [Fig. 6(a) and 6(b)]. These estimates are close to the numerically simulated in-focus lateral resolutions of  $8.9 \mu\text{m}$  (high NA) and  $18.1 \mu\text{m}$  (low NA). It is also noteworthy that the lateral resolution of the high-NA objective shows significant depth dependency. This can be accounted for by the short depth of focus ( $88.7 \mu\text{m}$  in theory and  $95.0 \mu\text{m}$  by Zemax simulation, in air), whereas the depth of focus of the low-NA objective is relatively long ( $355.0 \mu\text{m}$  in theory and  $392.8 \mu\text{m}$  by Zemax simulation, in air).

The average axial resolution estimates are  $14.1 \mu\text{m}$  and  $13.9 \mu\text{m}$ , respectively for the high- and low-NA objectives [Fig. 6(c) and 6(d)]. These estimates are close to the true depth resolution of  $14.1 \mu\text{m}$ .

### 3.3. Time-course evaluation of a tumor spheroid

The scatterer density estimator is applied to the time-lapse images of an *in vitro* human breast cancer spheroid measured up to 28 hours after the extraction from the cultivation environment (Fig. 7). The first and second rows [Fig. 7(a) and (b)] show the conventional OCT and scatterer density image, respectively. It is noted that the estimation window is relatively large, i.e.,  $31.2 \mu\text{m}$  (lateral)  $\times$   $115.8 \mu\text{m}$  (depth), in comparison with the spheroid size (around  $400 \mu\text{m}$  in diameter),

and the estimates at the spheroid periphery are thus not exactly reliable. Figures 7(c) and 7(d) show the tissue dynamics images [36] for reference. Figure 7(c) presents the “logarithmic intensity variance” (LIV), which is primarily sensitive to the magnitude of the signal intensity fluctuation, i.e., it is sensitive to the magnitude of the intracellular motility. Figure 7(d) presents the “OCT correlation decay speed at early delay time” ( $\text{OCDS}_e$ ), which reflects the speed of the signal decorrelation.  $\text{OCDS}_e$  has a large value if there is a rapid motion in the cell. These images are reprinted from our previous publication [36].

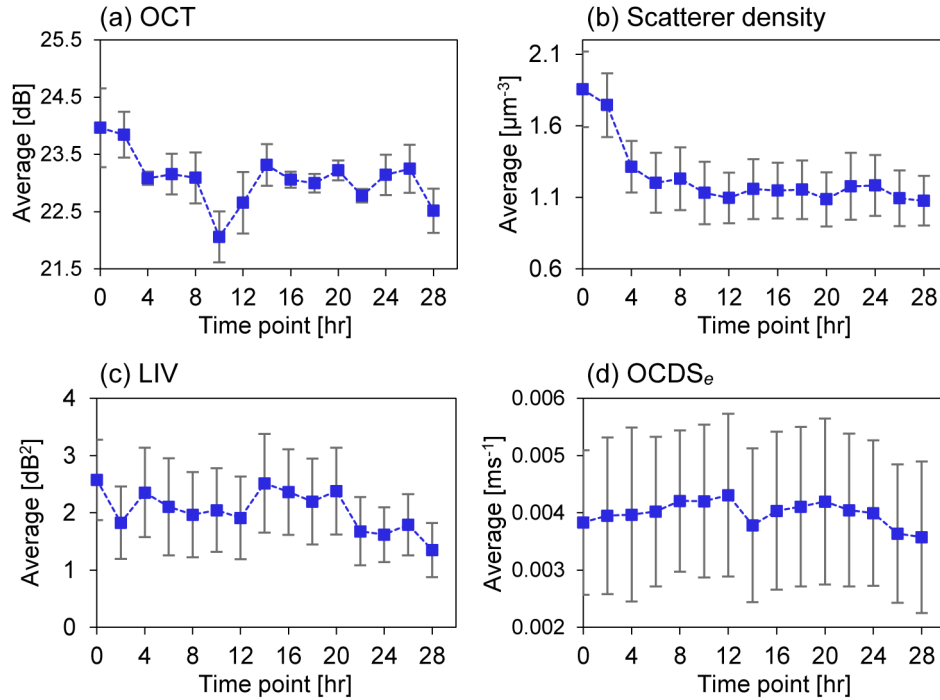


**Fig. 7.** Time-lapse images of a human breast adenocarcinoma spheroid (MCF7): (a) conventional OCT, (b) scatterer density images obtained using the DCNN-based scatterer density estimator, (c) magnitude of the signal fluctuation (LIV), (d) decorrelation speed of the OCT signal ( $\text{OCDS}_e$ ).

Figure 8 shows the time courses of the average OCT intensity (a), scatterer density (b), LIV (c), and  $\text{OCDS}_e$  (d) in regions of interest (ROI) around the center of the spheroid. The size of the ROI is lateral  $117 \mu\text{m} \times$  axial  $144.8 \mu\text{m}$  ( $60 \times 20$  pixels). The bars in the plots indicate the standard deviations.

The conventional OCT intensity [Fig. 7(a)] does not change appreciably over 28 hours and the average OCT signal intensity [Fig. 8] fluctuates with time and does not have a clear tendency. Meanwhile, the region with high scatterer density (green) rapidly becomes small over the first 8 hours [Fig. 7(b)]. Additionally, the average scatterer density [Fig. 8(b)] rapidly reduces over the first 8 hours. The low LIV region (red) and high  $\text{OCDS}_e$  region (yellow to green) became large with time [Fig. 7(c) and (d)]. Additionally, the average LIV gradually reduces with time [Fig. 8(c)], but this reduction is slower than that of the scatterer density [Fig. 8(b)].

The low LIV and high  $\text{OCDS}_e$  are thought to indicate cell necrosis [36,40]. The main scatterers in the cell are cell nuclei and cell organelles, and the rapid reduction of the scatterer density therefore suggests the destruction of nuclei and organelles during necrosis.



**Fig. 8.** Time courses of the average OCT intensity (a), scatterer density (b), LIV (c), and OCDS<sub>e</sub> (d) in the central region of the spheroid (lateral 117  $\mu\text{m} \times$  axial 144.8  $\mu\text{m}$ ).

## 4. Discussion

### 4.1. Scatterer diameter of the Intralipid

The scatterers of the Intralipid are the particles of soybean oil and egg lecithin [33–35]. Assuming the scatterer particles are spherical, the scatterer density (the number of scatterers per unit volume) of the Intralipid dilution can be computed as

$$\text{Scatterer density} = \frac{\sigma}{4\pi/3 (d/2)^3}, \quad (3)$$

where  $\sigma$  is volume concentration of the scattering medium and  $d$  is the diameter of the scatterer particle. Using this equation, the diameter of scatterer particle can be computed from the slope of the relation between the scatterer density and  $\sigma$  as

$$d = \sqrt[3]{6/(\pi \text{ slope})}. \quad (4)$$

In our Intralipid phantom measurements (Section 3.2.1), the slope of the relation between the scatterer density and  $\sigma$  is found to be  $0.5533 \pm 0.0086 \mu\text{m}^{-3}/\%(v/v)$ , or equivalently  $55.33 \pm 0.86 \mu\text{m}^{-3}/(v/v \text{ ratio})$  (mean  $\pm$  standard deviation over three attenuations and two NAs). The diameter of the scatterer particle is thus estimated to be 325.6 nm. Previously reported scatterer particle diameters of Intralipid vary in the literature; e.g., 25 to 473.9 nm [35], 50 to 400 nm with 225.7-nm mean diameter [33], and 360 nm [34]. The diameter estimated in the present study is well within the range of reported diameters.

#### 4.2. Benefits of numerically synthesized training data set

The present DCNN based estimators were trained by a numerically synthesized data set. The numerical generation of the training data set has several advantages against experimentally collecting a training data set.

At first, a massive amount of training data can be obtained with negligible cost. In the present study, 1,280,000 2D speckle patterns extracted from 80,000 3D synthesized OCT volumes were used for training (Section 2.2.1).

Second, the training data can have high diversity of the ground-truth fundamental parameter values. For instance, all the 80,000 synthesized OCT volumes were generated with different fundamental parameters in the present study. It might make the DCNN-based estimator robust. If we achieve the same diversity of the ground-truth parameters with an experimental data set, 80,000 accurately fabricated phantoms are required, and it is not realistic.

Third, some of the fundamental parameters are not really controllable in the experiment. For example, the lateral and axial resolutions are defined by the OCT device. And using massive amounts of OCT devices to prepare the training data set with a variety of resolutions is not realistic.

Finally, it is noteworthy that some of the fundamental parameters are hard to be accurately controlled in the experiment. As pointed out by Kübler *et al.*, the same recipe of the phantom does not always give the same scattering property [41]. In addition, the resolutions are affected by the aberrations, dispersion, wavelength-dependent absorption, and so on. And hence, it is hard to control or know the accurate resolutions in the experiment.

#### 4.3. Rationality of the estimation of the signal-strength-independent scatterer density

The speckle patterns and OCT images are normalized by intensity before being input to the DCNN-based estimator, and the estimator therefore cannot use the overall signal strength to estimate the scatterer density. This is consistent with the results shown in Fig. 4, where the scatterer density estimates are not sensitive to the probe beam attenuation.

This signal-strength-independent estimation process is reasonable for the following reasons. First, Hillman *et al.* showed that the contrast of a local speckle pattern has a unique and monotonical relationship with the ENS, i.e., the number of scatterers per coherence volume that is an alias of the 3D resolution [2]. This suggests that the ENS can be estimated from the speckle contrast in principle.

In addition, if we know the size of the coherence volume, we can estimate the scatterer density (i.e., the number of scatterers per unit volume) from the ENS. The size of the coherence volume can be estimated if the axial and lateral resolutions are known. Kurokawa *et al.* showed that these resolutions can be estimated through the spatial correlation analysis of a local speckle pattern [3].

That is to say, we can estimate both the ENS and the coherence volume size from the local speckle pattern. Hence, estimation of the scatterer density without information of the overall signal strength is possible in principle.

In the present study, the DCNN-based estimators successfully estimate the ENS and the resolutions (Sections 3.2.1 and 3.2.2), which further supports the rationality of the estimation.

The forgoing discussion gives the idea that the scatterer density can be computed by combining the speckle contrast analysis of Hillman and the correlation analysis of Kurokawa. However, it is difficult to accurately compute the speckle contrast and correlation because of noise in the speckle pattern. One strength of our DCNN-based estimator is that it can be trained to be robust against noise by adding random noise to the speckle patterns of the training data set.

#### 4.4. Limitation of the current estimators

The current estimation framework has limitations. First, the speckle generator does not consider the size of the scatterers. The speckle generator uses a 3D numerical field in which the pixel size

is  $0.24\ \mu\text{m} \times 0.24\ \mu\text{m} \times 0.90\ \mu\text{m}$  (lateral  $\times$  lateral  $\times$  depth), and a single scatterer correspond to a single pixel. This limited accuracy of the scatterer representation can reduce the accuracy of the estimation.

The second limitation is that the speckle generator does not account for depth-dependent signal attenuation. That is to say, the signals in the deeper region are not affected by the scattering or absorption in the superior region. Although we expect this effect to be small as the depth of the generated speckle pattern is only  $115.8\ \mu\text{m}$ , the estimator accuracy could be further improved by taking this effect into account.

Third, the sensitivity roll-off, depth-dependent variation of noise floor, and confocal effect [42] were not accounted for in our speckle generator. These factors can affect the estimation accuracy as it alters the SNR within the estimation window. However, for our small estimation window whose depth size is  $115.8\ \mu\text{m}$ , the effects of these factors might be limited. The estimation accuracy can be further improved by accounting these factors in the speckle generation process.

Fourth, in the speckle pattern generator, the scatterers are assumed to be static. Therefore, the jiggle of the Intralipid scatters due to Brownian motion is not accounted for. In addition to the above two limitations, this discrepancy would partially explain the non-zero intercept in Fig. 4.

Fifth, our study only considers fully developed speckle patterns and does not account for partially developed speckles. Although this might limit the accuracy and the applicability of the estimators, we believe our current estimators are still reasonable. It is because several previous studies of OCT are based on the assumption of fully developed speckles, and they have given reasonable results [1,2,43].

We expect further improvement of the speckle generator will increase the accuracy of the estimators.

#### 4.5. Effects of wavelength dispersion and polarization mode dispersion

In addition to the factors discussed in Section 4.4, wavelength dispersion and polarization mode dispersion (PMD) are also not considered in our speckle generator. It is known that the wavelength dispersion can be corrected numerically with image based metrics [44,45]. So, we can minimize its undesirable effect by applying the numerical dispersion correction to the OCT image before it is input to the estimators.

The PMD is known to deteriorate the OCT signal if the device is equipped with a very long optical fiber and/or improper optical components such as some types of circulators. And its undesirable effect is significant particularly for endoscopic polarization sensitive (PS-) OCT [46]. On the other hand, its effect is not significant for non-PS-OCT and non-endoscopic PS-OCT as far as correct optical components are used. And hence, the present study might not be significantly affected by the PMD. In order to apply our estimators to OCT with significant PMD, one of the several PMD correction methods [47–50] can be applied prior to the estimation.

## 5. Conclusion

We demonstrated DCNN-based estimators for the scatterer density, ENS, lateral and axial resolution, and SNR. The DCNN is trained using fully numerically generated OCT images (i.e., speckle patterns) and was therefore trained easily using an extremely large training data set comprising 1,280,000 images. Numerical validation showed the good performance of the estimators. Additionally, validation with the scattering phantoms showed the feasibility of the estimators in experiments. The scatterer density estimator was also applied to an *in vitro* tumor spheroid, and the reduction of the scatterer density during cell necrosis was visualized.

In conclusion, the DCNN-based estimators successfully extract fundamental parameters from a local speckle pattern. The scatterer density estimator can be used to quantify a cell microstructure smaller than the OCT resolution. Furthermore, the resolution estimators can be used as optimization metrics of image-based adaptive optics or computational aberration



correction and also for the calibration of OCT systems, such as in highly accurate measurements of the attenuation coefficient.

**Funding.** Japan Science and Technology Agency (JPMJCR2105, JPMJMI18G8); Japan Society for the Promotion of Science (18H01893, 21H01836); School of Science, King Mongkut's Institute of Technology Ladkrabang (RA/TA-2562-M-006).

**Acknowledgment.** The authors greatly appreciate the fruitful technical discussions that were held with Cheng-Kuang Lee from Nvidia.

**Disclosures.** Seesan, Abd El-Sadek, Mukherjee, Zhu, Oikawa: Yokogawa Electric Corp. (F), Sky technology (F), Nikon (F), Kao Corp. (F), Topcon (F). Miyazawa: Yokogawa Electric Corp. (F), Sky technology (F, E), Nikon (F), Kao Corp. (F), Topcon (F). Shen, Matsusaka: None. Makita, Yasuno: Yokogawa Electric Corp. (F), Sky technology (F), Nikon (F), Kao Corp. (F), Topcon (F), Tomey Corp. (P). Buranasiri: None.

**Data availability.** Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

## References

1. J. M. Schmitt, S. H. Xiang, and K. M. Yung, "Speckle in optical coherence tomography," *J. Biomed. Opt.* **4**(1), 95–105 (1999).
2. T. R. Hillman, S. G. Adie, V. Seemann, J. J. Armstrong, S. L. Jacques, and D. D. Sampson, "Correlation of static speckle with sample properties in optical coherence tomography," *Opt. Lett.* **31**(2), 190–192 (2006).
3. K. Kurokawa, S. Makita, Y.-J. Hong, and Y. Yasuno, "In-plane and out-of-plane tissue micro-displacement measurement by correlation coefficients of optical coherence tomography," *Opt. Lett.* **40**(9), 2153–2156 (2015).
4. P. Wijesinghe, L. Chin, and B. F. Kennedy, "Strain tensor imaging in compression optical coherence elastography," *IEEE J. Sel. Topics Quantum Electron.* **25**(1), 1–12 (2019).
5. K. A. Vermeer, J. van der Schoot, H. G. Lemij, and J. F. de Boer, "RPE-normalized RNFL attenuation coefficient maps derived from volumetric OCT imaging for glaucoma assessment," *Invest. Ophthalmol. Visual Sci.* **53**(10), 6102–6108 (2012).
6. K. A. Vermeer, J. Mo, J. J. A. Weda, H. G. Lemij, and J. F. de Boer, "Depth-resolved model-based reconstruction of attenuation coefficients in optical coherence tomography," *Biomed. Opt. Express* **5**(1), 322–337 (2014).
7. T. M. Cannon, B. E. Bouma, and N. Uribe-Patarroyo, "Layer-based, depth-resolved computation of attenuation coefficients and backscattering fractions in tissue using optical coherence tomography," *Biomed. Opt. Express* **12**(8), 5037–5056 (2021).
8. A. C. Chan, K. Kurokawa, S. Makita, M. Miura, and Y. Yasuno, "Maximum a posteriori estimator for high-contrast image composition of optical coherence tomography," *Opt. Lett.* **41**(2), 321–324 (2016).
9. A. C. Chan, Y.-J. Hong, S. Makita, M. Miura, and Y. Yasuno, "Noise-bias and polarization-artifact corrected optical coherence tomography by maximum a-posteriori intensity estimation," *Biomed. Opt. Express* **8**(4), 2069–2087 (2017).
10. D. Kasaragod, S. Makita, S. Fukuda, S. Beheregaray, T. Oshika, and Y. Yasuno, "Bayesian maximum likelihood estimator of phase retardation for quantitative polarization-sensitive optical coherence tomography," *Opt. Express* **22**(13), 16472–16492 (2014).
11. S. Makita, Y.-J. Hong, M. Miura, and Y. Yasuno, "Degree of polarization uniformity with high noise immunity using polarization-sensitive optical coherence tomography," *Opt. Lett.* **39**(24), 6783–6786 (2014).
12. N. Lippok, M. Villiger, and B. E. Bouma, "Degree of polarization uniformity and depolarization index: unambiguous depolarization contrast for optical coherence tomography," *Opt. Lett.* **40**(17), 3954–3957 (2015).
13. M. Yamanari, S. Tsuda, T. Kokubun, K. Omodaka, N. Aizawa, Y. Yokoyama, S. Kunimatsu-Sanuki, K. Maruyama, and T. Nakazawa, "Estimation of Jones matrix, birefringence and entropy using Cloude-Pottier decomposition in polarization-sensitive optical coherence tomography," *Biomed. Opt. Express* **7**(9), 3551–3573 (2016).
14. D. Kasaragod, S. Makita, Y.-J. Hong, and Y. Yasuno, "Noise stochastic corrected maximum a posteriori estimator for birefringence imaging using polarization-sensitive optical coherence tomography," *Biomed. Opt. Express* **8**(2), 653–669 (2017).
15. L. Duan, S. Makita, M. Yamanari, Y. Lim, and Y. Yasuno, "Monte-Carlo-based phase retardation estimator for polarization sensitive optical coherence tomography," *Opt. Express* **19**(17), 16330–16345 (2011).
16. L. Chin, A. Curatolo, B. F. Kennedy, B. J. Doyle, P. R. T. Munro, R. A. McLaughlin, and D. D. Sampson, "Analysis of image formation in optical coherence elastography using a multiphysics approach," *Biomed. Opt. Express* **5**(9), 2913–2930 (2014).
17. P. R. Munro, A. Curatolo, and D. D. Sampson, "Full wave model of image formation in optical coherence tomography applicable to general samples," *Opt. Express* **23**(3), 2541–2556 (2015).
18. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* **521**(7553), 436–444 (2015).
19. M. M. Najafabadi, F. Villanustre, T. M. Khoshgoftaar, N. Seliya, R. Wald, and E. Muharemagic, "Deep learning applications and challenges in big data analytics," *J. Big Data* **2**(1), 1 (2015).



20. X. Sui, Y. Zheng, B. Wei, H. Bi, J. Wu, X. Pan, Y. Yin, and S. Zhang, "Choroid segmentation from optical coherence tomography with graph-edge weights learned from deep convolutional neural networks," *Neurocomputing* **237**, 332–341 (2017).
21. F. G. Venhuizen, B. van Ginneken, B. Liefers, M. J. van Grinsven, S. Fauser, C. Hoyng, T. Theelen, and C. I. Sánchez, "Robust total retina thickness segmentation in optical coherence tomography images using convolutional neural networks," *Biomed. Opt. Express* **8**(7), 3292–3316 (2017).
22. A. G. Roy, S. Conjeti, S. P. K. Karri, D. Sheet, A. Katouzian, C. Wachinger, and N. Navab, "ReLayNet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks," *Biomed. Opt. Express* **8**(8), 3627–3642 (2017).
23. M. Pekala, N. Joshi, T. A. Liu, N. Bressler, D. C. DeBuc, and P. Burlina, "Deep learning based retinal OCT segmentation," *Comput. Biol. Med.* **114**, 103445 (2019).
24. C. S. Lee, A. J. Tying, N. P. Deruyter, Y. Wu, A. Rokem, and A. Y. Lee, "Deep-learning based, automated segmentation of macular edema in optical coherence tomography," *Biomed. Opt. Express* **8**(7), 3440–3448 (2017). Publisher: Optical Society of America.
25. H. Mohsen, E.-S. A. El-Dahshan, E.-S. M. El-Horbaty, and A.-B. M. Salem, "Classification using deep learning neural networks for brain tumors," *Futur. Comput. Inform. J.* **3**(1), 68–71 (2018).
26. Y. Ma, X. Chen, W. Zhu, X. Cheng, D. Xiang, and F. Shi, "Speckle noise reduction in optical coherence tomography images based on edge-sensitive cGAN," *Biomed. Opt. Express* **9**(11), 5129–5146 (2018).
27. S. K. Devalla, G. Subramanian, T. H. Pham, X. Wang, S. Perera, T. A. Tun, T. Aung, L. Schmetterer, A. H. Thiéry, and M. J. A. Girard, "A deep learning approach to denoise optical coherence tomography images of the optic nerve head," *Sci. Rep.* **9**(1), 14454 (2019).
28. G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. van der Laak, B. van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Med. Image Anal.* **42**, 60–88 (2017).
29. R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," *Insights Imag.* **9**(4), 611–629 (2018).
30. A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, "A survey of the recent architectures of deep convolutional neural networks," *Artif. Intell. Rev.* **53**(8), 5455–5516 (2020).
31. D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," arXiv:1412.6980 [cs] (2017). ArXiv: 1412.6980.
32. B. W. Pogue and M. S. Patterson, "Review of tissue simulating phantoms for optical spectroscopy, imaging and dosimetry," *J. Biomed. Opt.* **11**(4), 041102 (2006).
33. X. Wen, V. V. Tuchin, Q. Luo, and D. Zhu, "Controlling the scattering of Intralipid by using optical clearing agents," *Phys. Med. Biol.* **54**(22), 6917–6930 (2009).
34. G. Zaccanti, S. Del Bianco, and F. Martelli, "Measurements of optical properties of high-density media," *Appl. Opt.* **42**(19), 4023–4030 (2003).
35. R. Michels, F. Foschum, and A. Kienle, "Optical properties of fat emulsions," *Opt. Express* **16**(8), 5907–5925 (2008).
36. I. Abd El-Sadek, A. Miyazawa, L. Tzu-Wei Shen, S. Makita, S. Fukuda, T. Yamashita, Y. Oka, P. Mukherjee, S. Matsusaka, T. Oshika, H. Kano, and Y. Yasuno, "Optical coherence tomography-based tissue dynamics imaging for longitudinal and drug response evaluation of tumor spheroids," *Biomed. Opt. Express* **11**(11), 6231–6248 (2020).
37. E. Li, S. Makita, Y.-J. Hong, D. Kasaragod, and Y. Yasuno, "Three-dimensional multi-contrast imaging of in vivo human skin by Jones matrix optical coherence tomography," *Biomed. Opt. Express* **8**(3), 1290–1305 (2017).
38. M. J. Ju, Y.-J. Hong, S. Makita, Y. Lim, K. Kurokawa, L. Duan, M. Miura, S. Tang, and Y. Yasuno, "Advanced multi-contrast Jones matrix optical coherence tomography for Doppler and polarization sensitive imaging," *Opt. Express* **21**(16), 19412–19436 (2013).
39. J. G. Fujimoto and W. Drexler, "Introduction to OCT," in *Optical Coherence Tomography: Technology and Applications*, W. Drexler and J. G. Fujimoto, eds. (Springer International Publishing, Switzerland, 2015), pp. 3–64.
40. G. Farhat, A. Mariampillai, V. X. D. Yang, G. J. Czarnota, and M. C. Kolios, "Optical coherence tomography speckle decorrelation for detecting cell death," *Proc. SPIE* **7907**, 790710 (2011).
41. J. Kübler, V. S. Zoutenbier, A. Amelink, J. Fischer, and J. F. de Boer, "Investigation of methods to extract confocal function parameters for the depth resolved determination of attenuation coefficients using OCT in intralipid samples, titanium oxide phantoms, and in vivo human retinas," *Biomed. Opt. Express* **12**(11), 6814–6830 (2021).
42. T. van Leeuwen, D. Faber, and M. Aalders, "Measurement of the axial point spread function in scattering media using single-mode fiber-based optical coherence tomography," *IEEE J. Sel. Topics Quantum Electron.* **9**(2), 227–233 (2003).
43. B. Karamata, K. Hassler, M. Laubscher, and T. Lasser, "Speckle statistics in optical coherence tomography," *J. Opt. Soc. Am. A* **22**(4), 593–596 (2005).
44. M. Wojtkowski, V. J. Srinivasan, T. H. Ko, J. G. Fujimoto, A. Kowalczyk, and J. S. Duker, "Ultrahigh-resolution, high-speed, Fourier domain optical coherence tomography and methods for dispersion compensation," *Opt. Express* **12**(11), 2404–2422 (2004).
45. Y. Yasuno, Y. Hong, S. Makita, M. Yamanari, M. Akiba, M. Miura, and T. Yatagai, "In vivo high-contrast imaging of deep posterior eye by 1- $\mu$ m swept source optical coherence tomography and scattering optical coherence angiography," *Opt. Express* **15**(10), 6121–6139 (2007).

46. M. Villiger, E. Z. Zhang, S. Nadkarni, W.-Y. Oh, B. E. Bouma, and B. J. Vakoc, "Artifacts in polarization-sensitive optical coherence tomography caused by polarization mode dispersion," *Opt. Lett.* **38**(6), 923–925 (2013).
47. E. Z. Zhang, W.-Y. Oh, M. L. Villiger, L. Chen, B. E. Bouma, and B. J. Vakoc, "Numerical compensation of system polarization mode dispersion in polarization-sensitive optical coherence tomography," *Opt. Express* **21**(1), 1163–1180 (2013).
48. M. Villiger, E. Z. Zhang, S. K. Nadkarni, W.-Y. Oh, B. J. Vakoc, and B. E. Bouma, "Spectral binning for mitigation of polarization mode dispersion artifacts in catheter-based optical frequency domain imaging," *Opt. Express* **21**(14), 16353–16369 (2013).
49. B. Braaf, K. A. Vermeer, M. de Groot, K. V. Vienola, and J. F. de Boer, "Fiber-based polarization-sensitive OCT of the human retina with correction of system polarization distortions," *Biomed. Opt. Express* **5**(8), 2736–2758 (2014).
50. Y.-J. Hong, S. Makita, S. Sugiyama, and Y. Yasuno, "Optically buffered Jones-matrix-based multifunctional optical coherence tomography with polarization mode dispersion correction," *Biomed. Opt. Express* **6**(1), 225–243 (2015).